第24回 蛋白質科学会年会

WS19 モデル生物大腸菌における翻訳制御研究の最前線

Reconstituted cell-free protein synthesis kit
PURE*frex*®

大腸菌由来の再構成型無細胞系内での翻訳反応は大腸菌細胞と同一か?

金森 崇、布施(村上) 朋重、松本 令奈（ジーンフロンティア(株)）

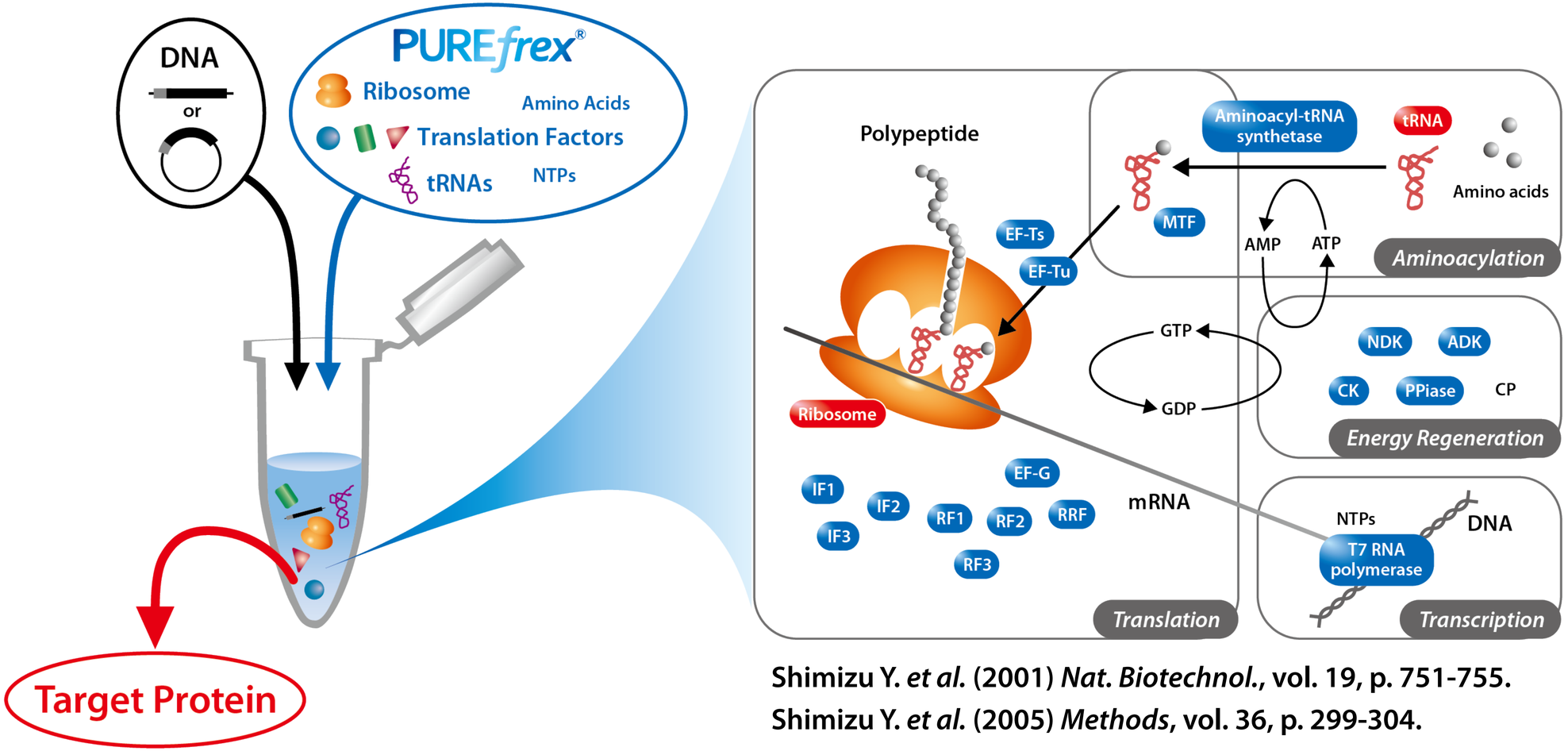
2024. 6. 13

GeneFrontier

PURE system

PURE*frex* is based on the PURE system technology.

The PURE system is a reconstituted cell-free protein synthesis system, which consists only of purified factors necessary for transcription, translation and energy regeneration in *Escherichia coli*.



Shimizu Y. *et al.* (2001) *Nat. Biotechnol.*, vol. 19, p. 751-755.

Shimizu Y. *et al.* (2005) *Methods*, vol. 36, p. 299-304.

GeneFrontier

Topics

- ★ PURE system (Translation in *E. coli*)
- ★ Initiation in the PURE system
- ★ Optimum sequence for the PURE system
  - ▸ 5’-UTR
  - ▸ N-terminal region in the ORF

GeneFrontier

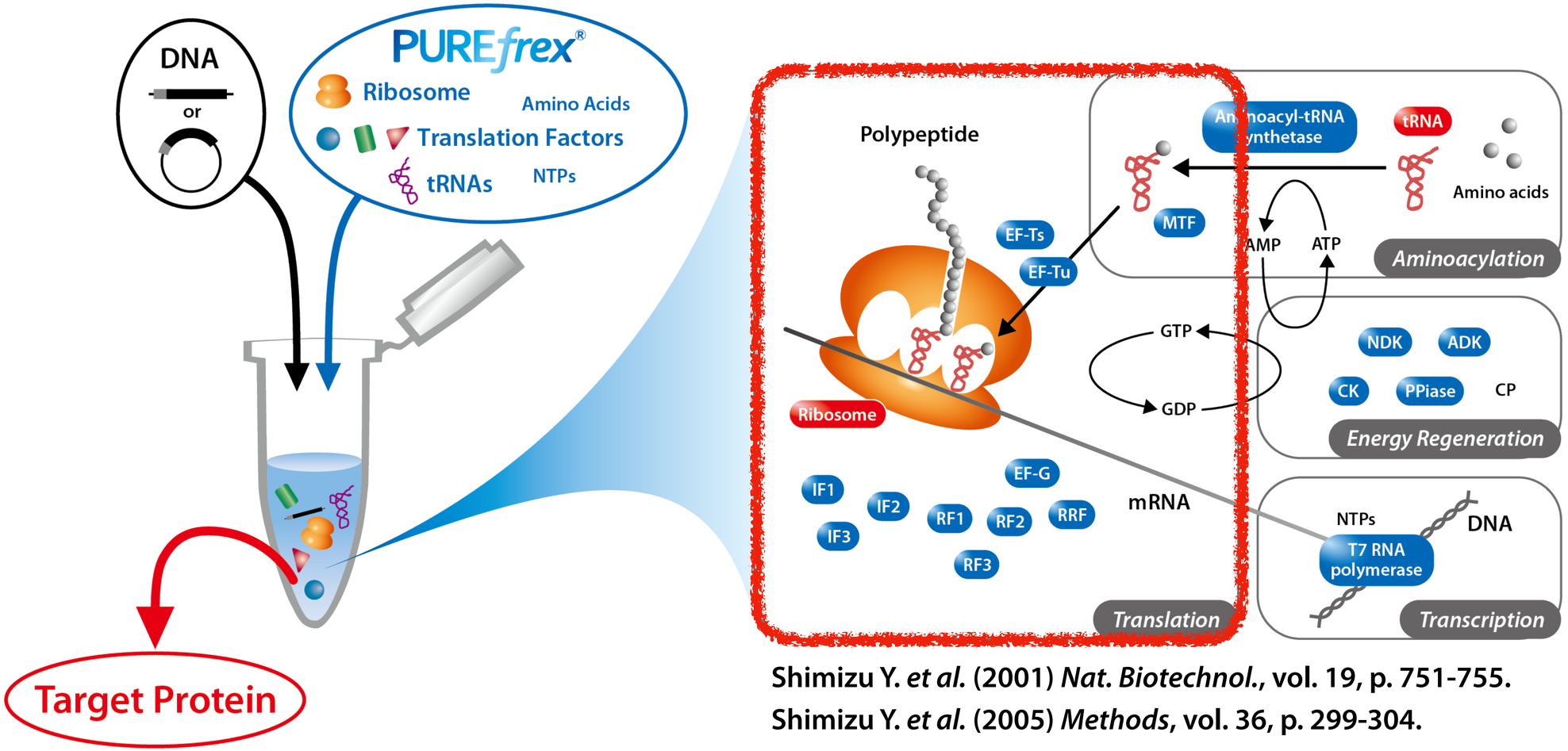
Topics

- ★ **PURE system (Translation in *E. coli*)**
- ★ Initiation in the PURE system
- ★ Optimum sequence for the PURE system
  - ▸ 5’-UTR
  - ▸ N-terminal region in the ORF

PURE system

PURE*frex* is based on the PURE system technology.

The PURE system is a reconstituted cell-free protein synthesis system, which consists only of purified factors necessary for transcription, translation and energy regeneration in *Escherichia coli*.

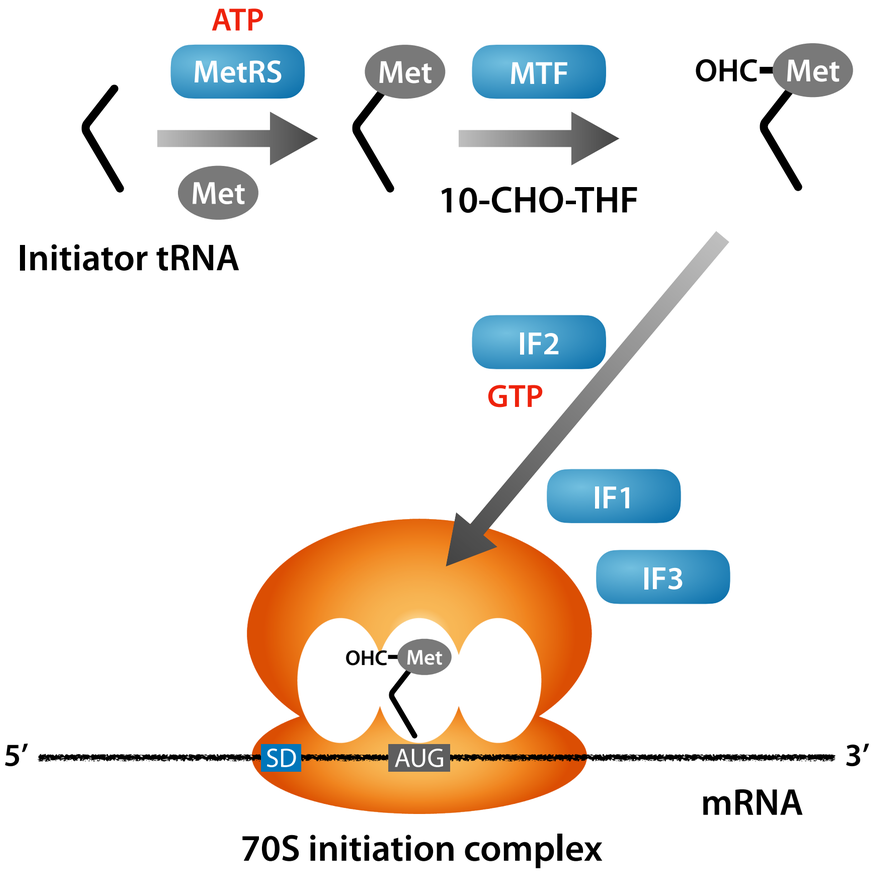


Shimizu Y. *et al.* (2001) *Nat. Biotechnol.*, vol. 19, p. 751-755.

Shimizu Y. *et al.* (2005) *Methods*, vol. 36, p. 299-304.

GeneFrontier

Initiation



MTF (Methionyl-tRNA transformylase)

- transfers formyl group from 10-formyl-tetrahydrofolic acid (10-CHO-THF) to the free amino group in methionine bound to initiator tRNA

IF1 (Initiation Factor 1)

- binds the A-site in the 30S subunit and prevents binding of aminoacyl-tRNA.
- modulates and stabillizes binding of IF2 and IF3 to the 30S subunit.

IF2 (Initiation Factor 2)

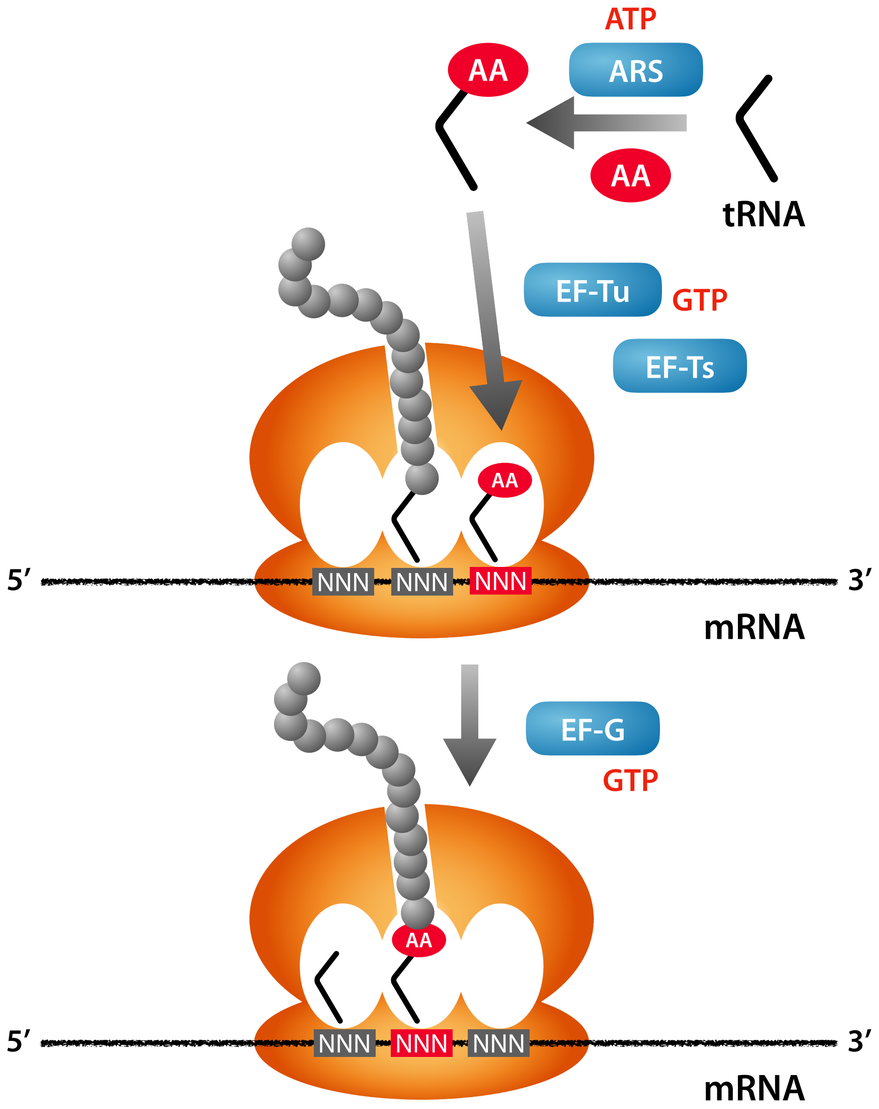
- is a GTPase and hydrolyzes GTP to GDP during formation of the 70S initiation complex.
- binds formylmethionyl-initiator tRNA and promotes it the entry to the P-site.
- is expressed as isoforms (α, β, β′) from in-frame different AUG codons of *infB* gene.

IF3 (Initiation Factor 3)

- binds to the 30S subunit and prevents association of the premature 50S subunit.
- enhances the accomodation of formylmethionyl-initiator tRNA to the P-site.
- Its gene contains the non-canonical initation codon AUU.

All of four proteins are essential for cell growth.

Elongation



EF-Tu (Elongation Factor Tu)

- delivers aminoacyl-tRNA to the A-site of ribosome in a GTP-boud form.
- is encoded by two genes, *tufA* and *tufB*.

EF-Ts (Elongation Factor Ts)

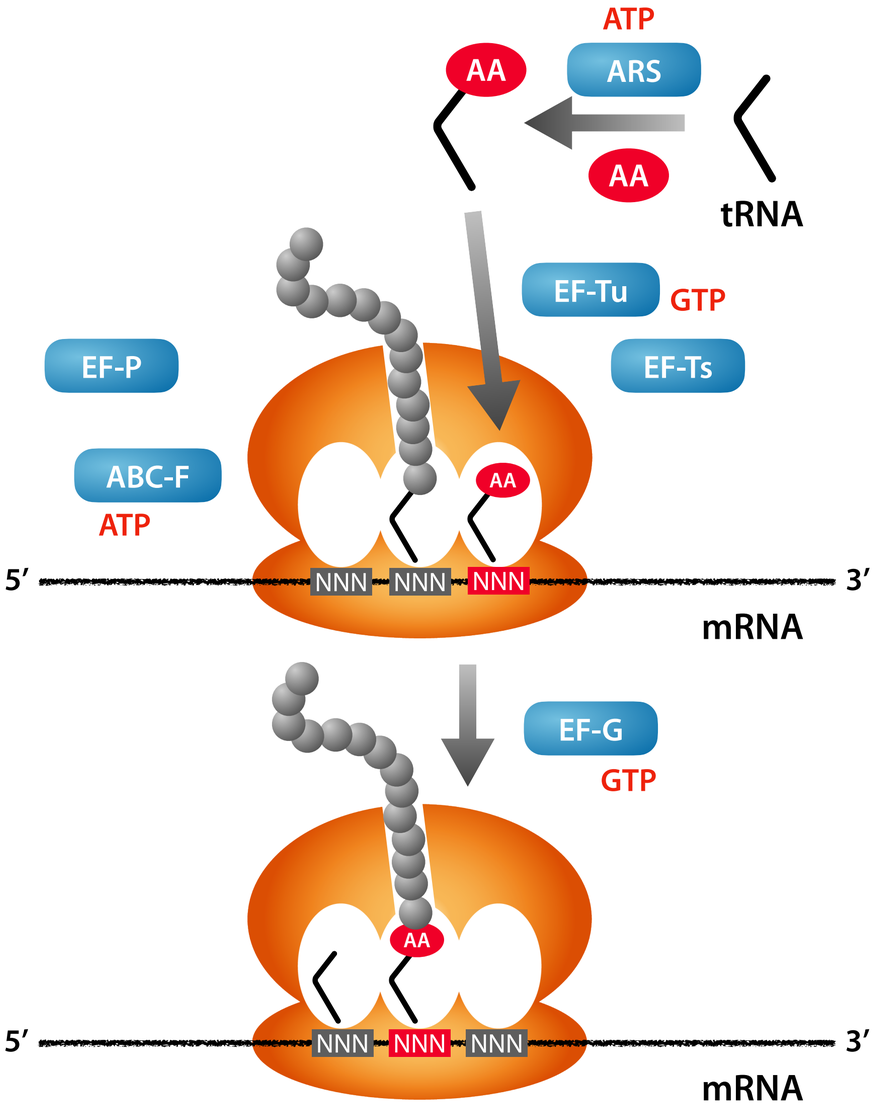
- associates GDP-bound EF-Tu and stimulate the exchange of GDP and GTP.

EF-G (Elongation Factor G)

- facislitates the translocation of the ribosome in a GTP hydrolysis-dependent manner.

All of three EFs are essential for cell growth.

Elongation



EF-Tu (Elongation Factor Tu)

- delivers aminoacyl-tRNA to the A-site of ribosome in a GTP-boud form.
- is encoded by two genes, *tufA* and *tufB*.

EF-Ts (Elongation Factor Ts)

- associates GDP-bound EF-Tu and stimulate the exchange of GDP and GTP.

EF-G (Elongation Factor G)

- facislitates the translocation of the ribosome in a GTP hydrolysis-dependent manner.

EF-P and ABC-F proteins

- promotes the formation of the peptide bond for the specific sequences.

  EF-P: poly-Pro

  YheS: SecM, Uup: poly-Pro, YbiT: K/E repeat and D/E repeat, EttA: D/E repeat

Ude *et al.* (2013) *Science*, 339, 82-85.

Doerfel *et al.* (2013) *Science*, 339, 85-88.

Murina *et al.* (2019) *J. Mol. Biol.*, 431, 3568-3590.

Chadani *et al.* (2024) *Nucleic Acids Res.*, gkae309.

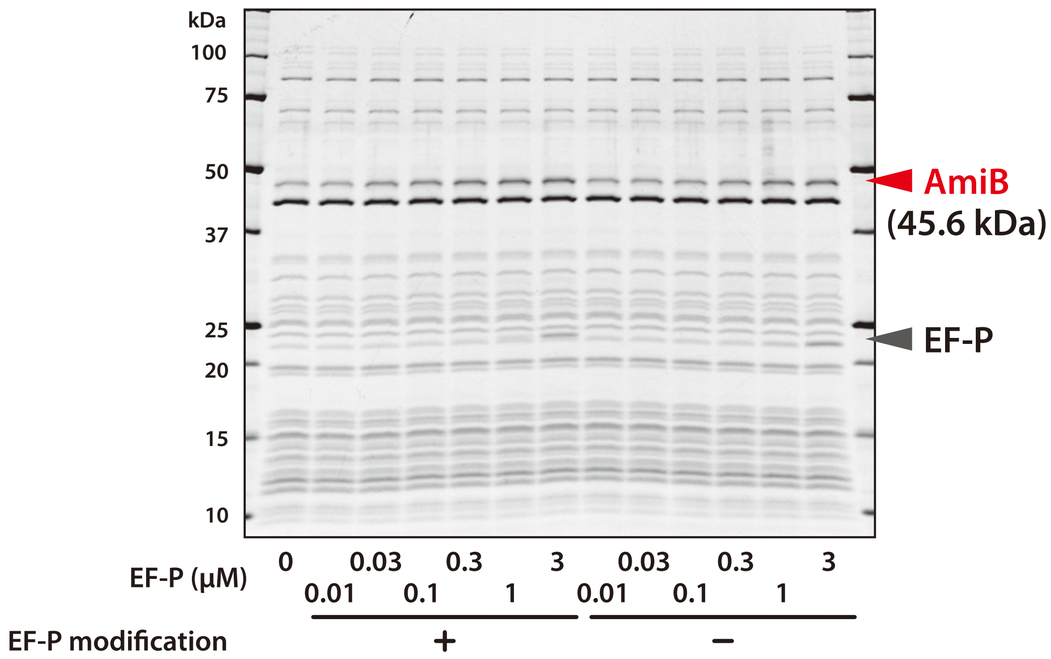
GeneFrontier

Elongation (Effect of EF-P)

AmiB (*E. coli*)

129

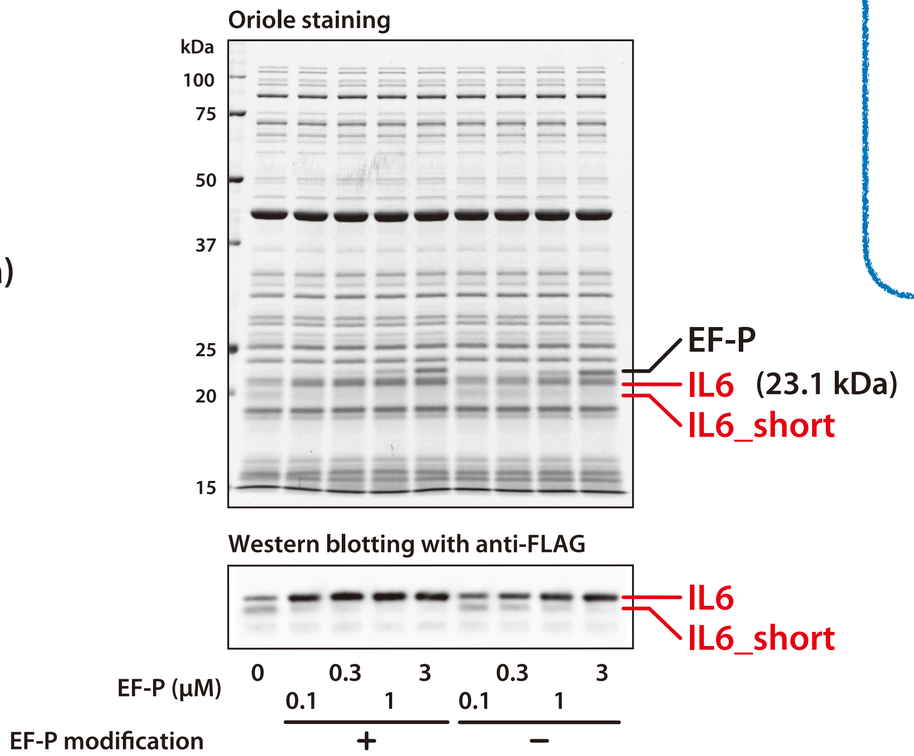
VPPPPPPPPPV

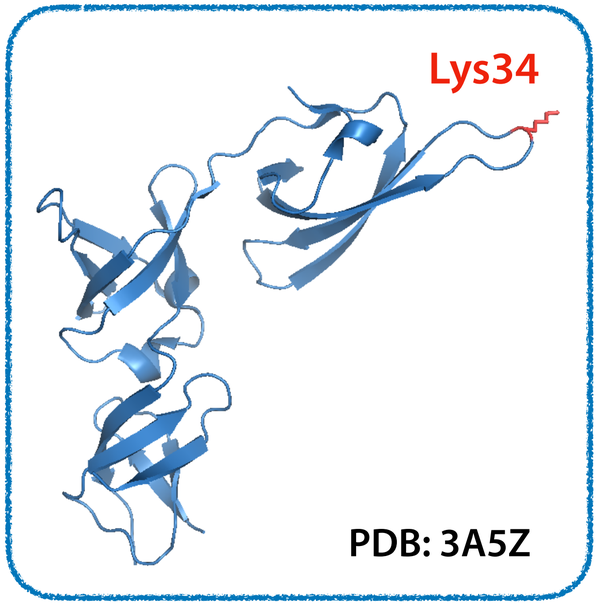


IL6 (mature) (*H. sapiens*)

1

MVPPGE

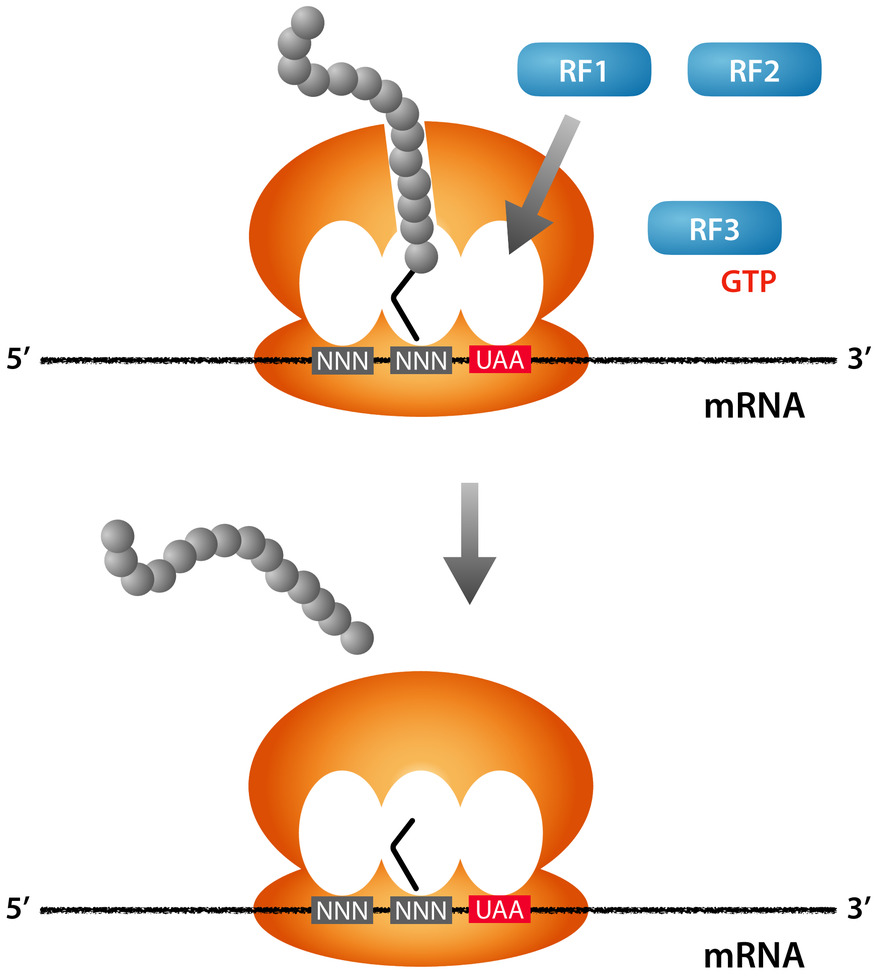




EF-P enhanced the synthesis of proteins with Pro-Pro sequence in the PURE system.

GeneFrontier

Termination



RF1 (Release Factor 1)

- recognizes the stop codon UAG and UAA, and releases the nascent polypeptide chain.
- is post-translationally methylated at the position Gln 235 in the conserved GGQ motif.

RF2 (Release Factor 2)

- recognizes the stop codon UGA and UAA, and releases the nascent polypeptide chain.
- is post-translationally methylated at the position Gln 252 in the conserved GGQ motif.
- Its gene contains UGA codon at the position 26 and the full-length protein is synthesized by +1 frameshft.
- Substitution of Thr at the position 246 to Ala increases the termination efficiency.

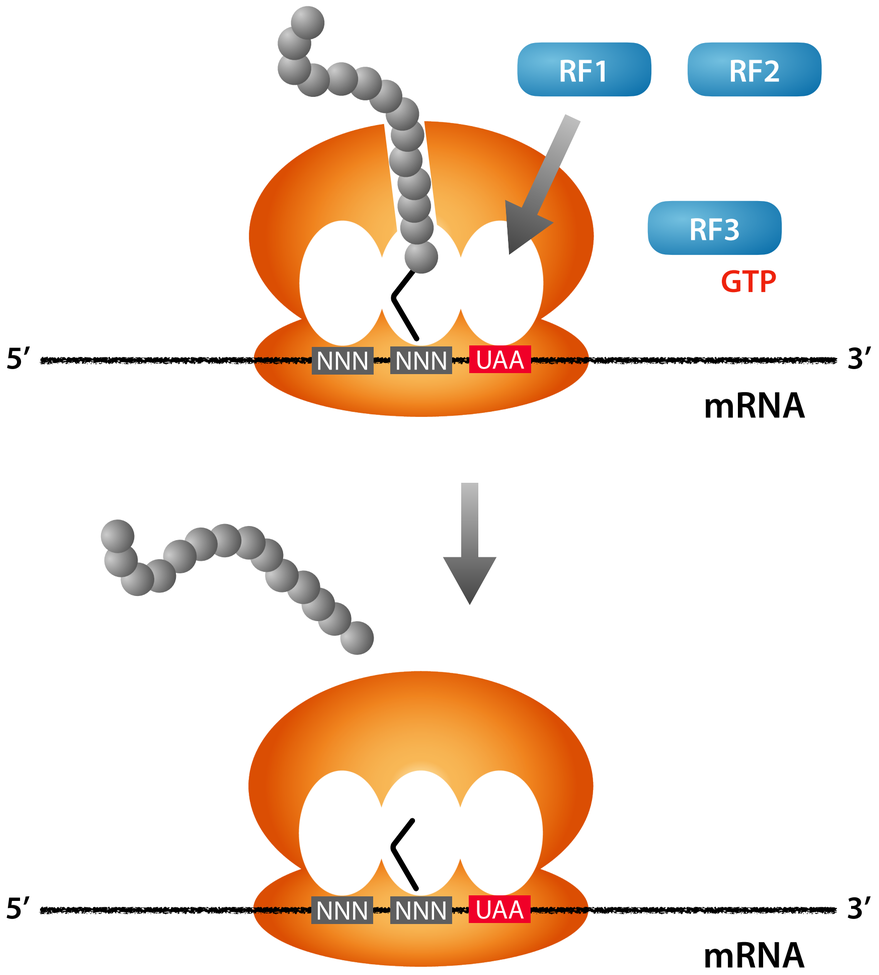
RF3 (Release Factor 3)

- is a ribosome-dependent GTPase.
- promotes the release of RF1 and RF2 from the ribosome after the release of the peptide chain.

RF1 and RF2 are essential for cell growth.

GeneFrontier

Termination



RF1 (Release Factor 1)

- recognizes the stop codon UAG and UAA, and releases the nascent polypeptide chain.
- is post-translationally methylated at the position Gln 235 in the conserved GGQ motif.

RF2 (Release Factor 2)

- recognizes the stop codon UGA and UAA, and releases the nascent polypeptide chain.
- is post-translationally methylated at the position Gln 252 in the conserved GGQ motif.
- Its gene contains UGA codon at the position 26 and the full-length protein is synthesized by +1 frameshft.
- Substitution of Thr at the position 246 to Ala increases the termination efficiency.

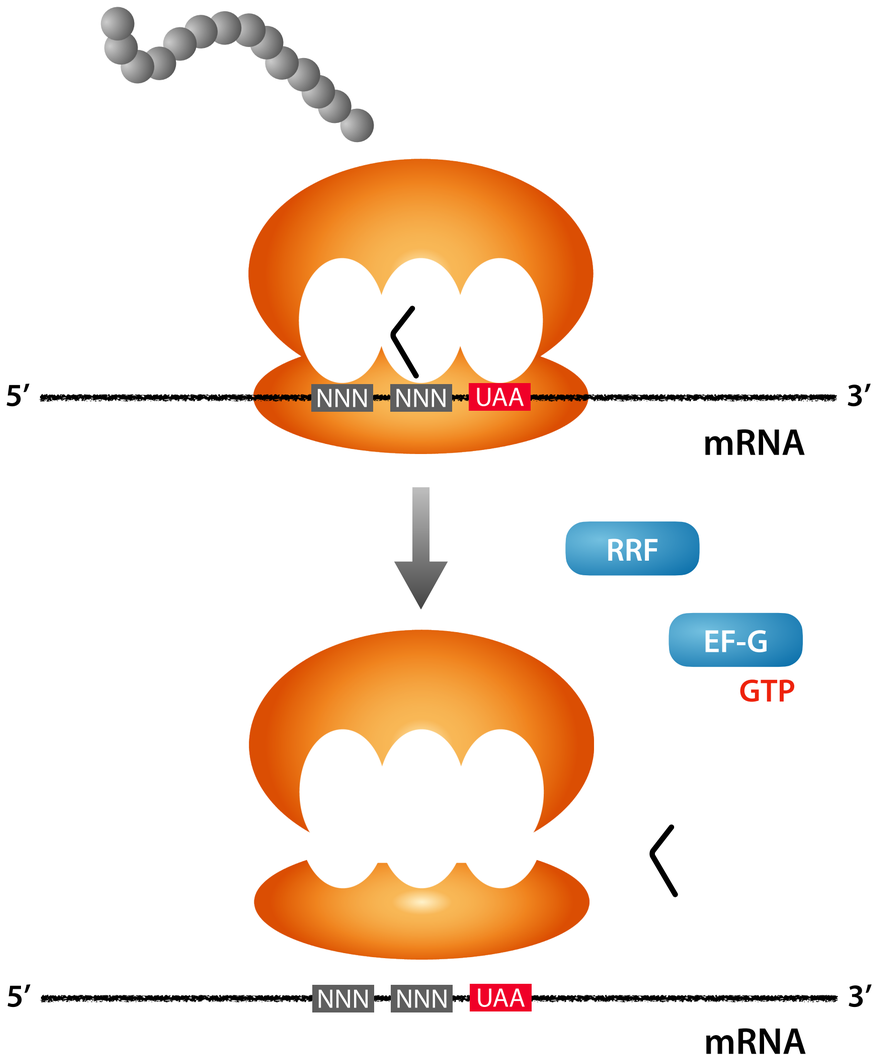
RF3 (Release Factor 3)

- is a ribosome-dependent GTPase.
- promotes the release of RF1 and RF2 from the ribosome after the release of the peptide chain.

RF1 and RF2 are essential for cell growth.

GeneFrontier

Recycling



RRF (Ribosome Recycling Factor)

- Is involved in ribosome recycling with EF-G after the termination process.

RRF is essential for cell growth.

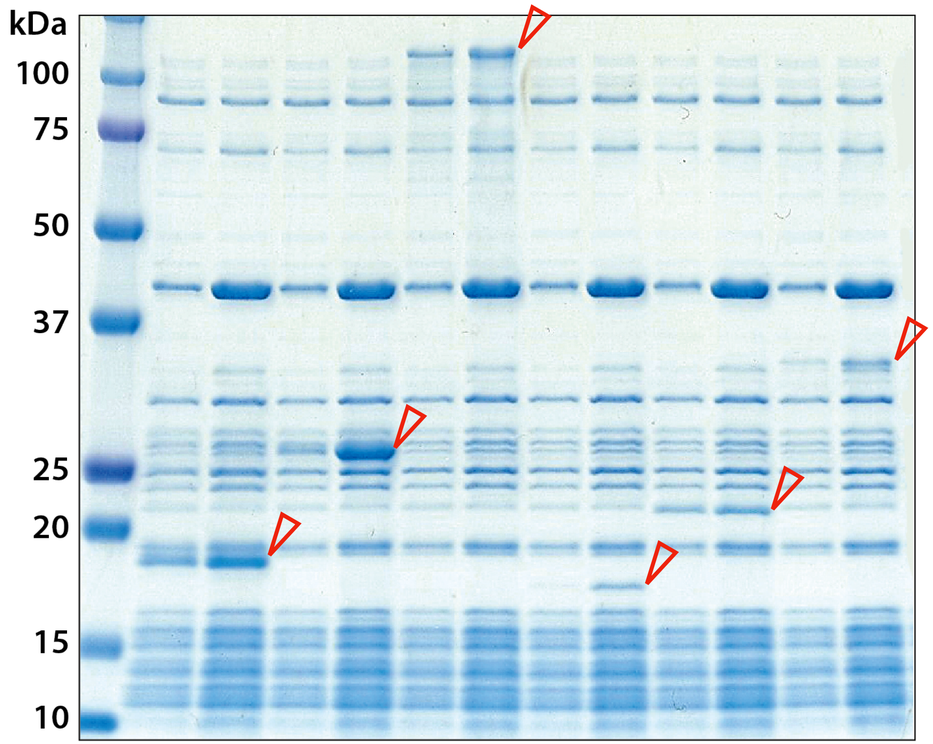
GeneFrontier

E. coli-based cell-free protein sysnthesis systems

	Extract system	Reconstituted system		
	S30 system	PURE system (original)	PURE*frex*®1.0	PURE*frex*®2.0
Typical Yield (μg/mL)	100-1,000	10-200	10-200	20-1,000
Contamination				
RNase	very High	Low	very Low	very Low
LPS	very High	High	very Low	very Low
Template DNA				
Plasmid DNA	OK	OK	OK	OK
PCR product	NG	OK	OK	OK
Customization of Reagent	Difficult	Easy	Easy	Easy
Purification of His-tagged product	OK	NG	OK	OK

Shimizu Y. *et al.* (2001) *Nat. Biotechnol.*, vol. 19, p. 751-755.
Shimizu Y. *et al.* (2005) *Methods*, vol. 36, p. 299-304.

Example of protein synthesis



DHFR:
Dihydrofolate reductase
GFP:
Green Fluorescent Protein
β-Gal:
β-Galactosidase
α-Syn:
α-Synuclein
bR:
Bacteriorhodopsin
MDH:
Malate Dehydrogenase

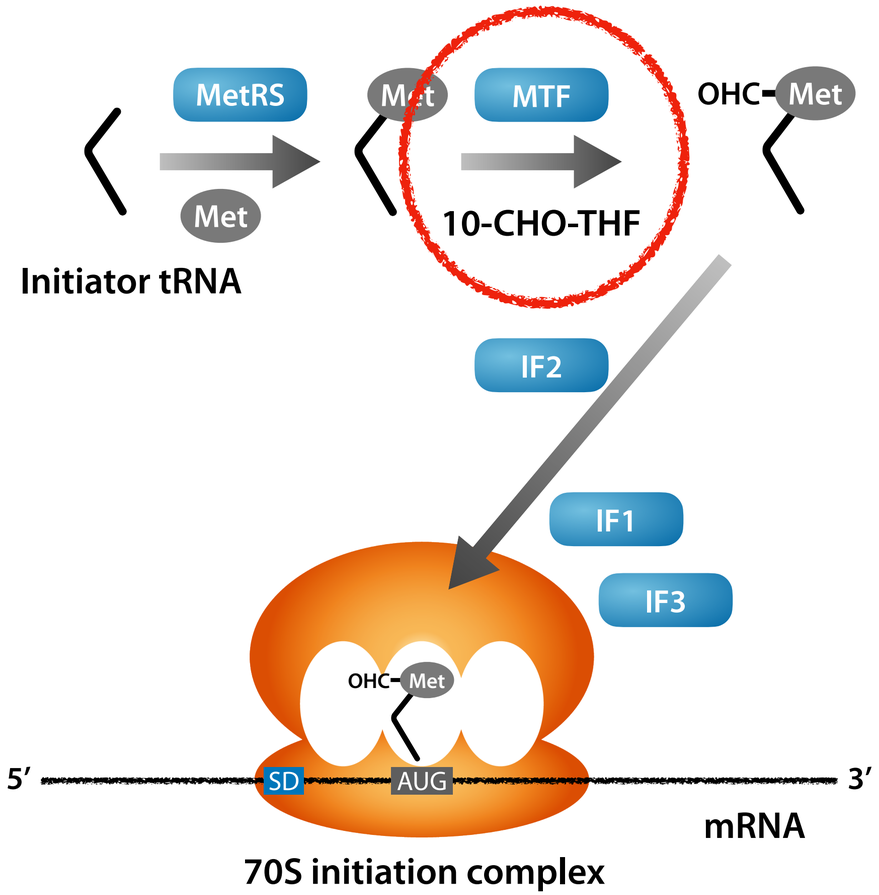
GeneFrontier

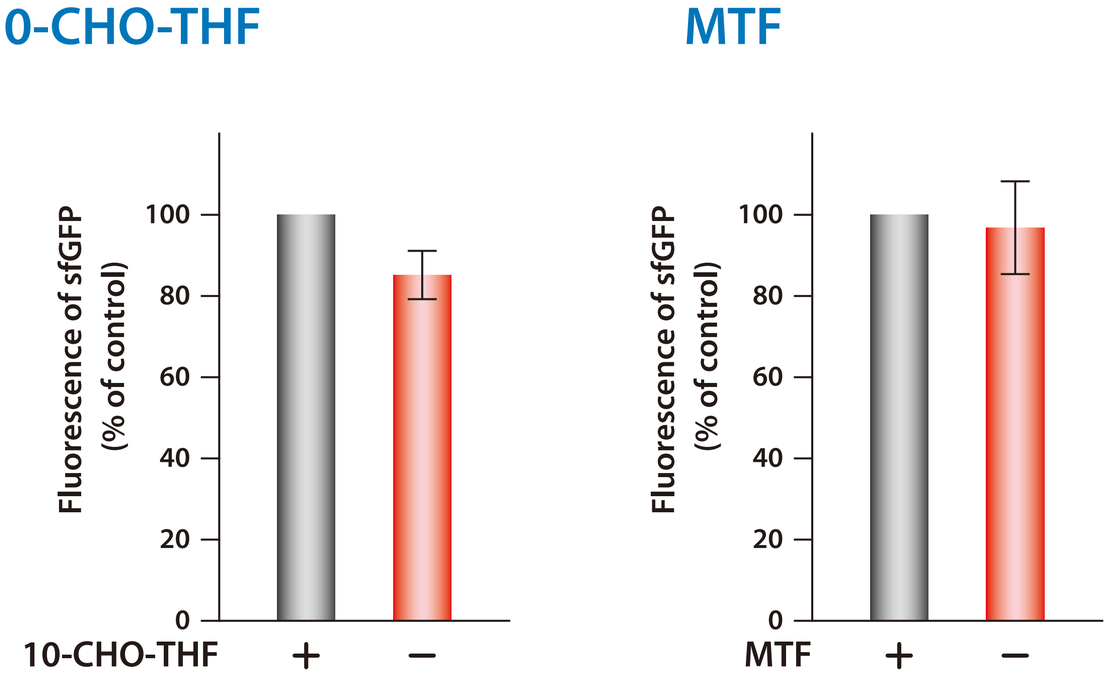
Topics

- ★ PURE system (Translation in *E. coli*)
- ★ **Initiation in the PURE system**
- ★ Optimum sequence for the PURE system
  - ▸ 5’-UTR
  - ▸ N-terminal region in the ORF

GeneFrontier

Initiation in the PURE system (MTF and 10 -CHO-THF)

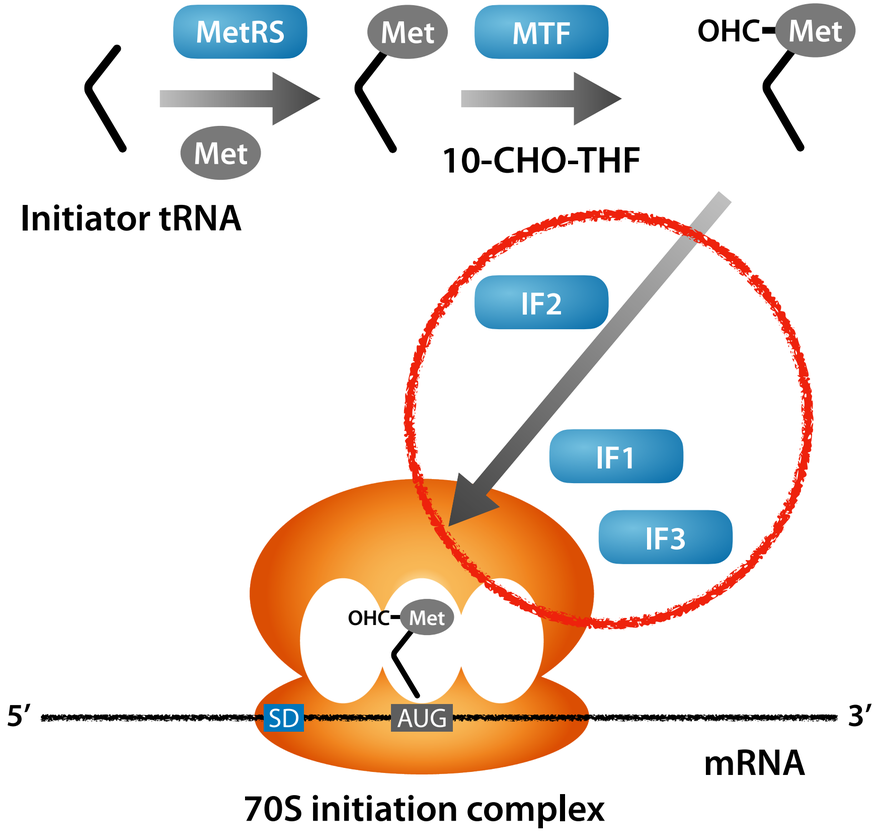




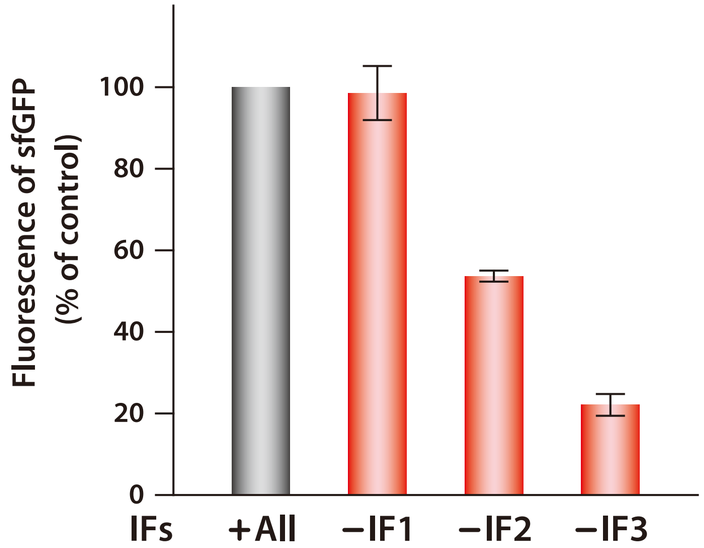
Formylation of initiator methionine was not essential for the translation in the PURE system.

GeneFrontier

Initiation in the PURE system (IFs)



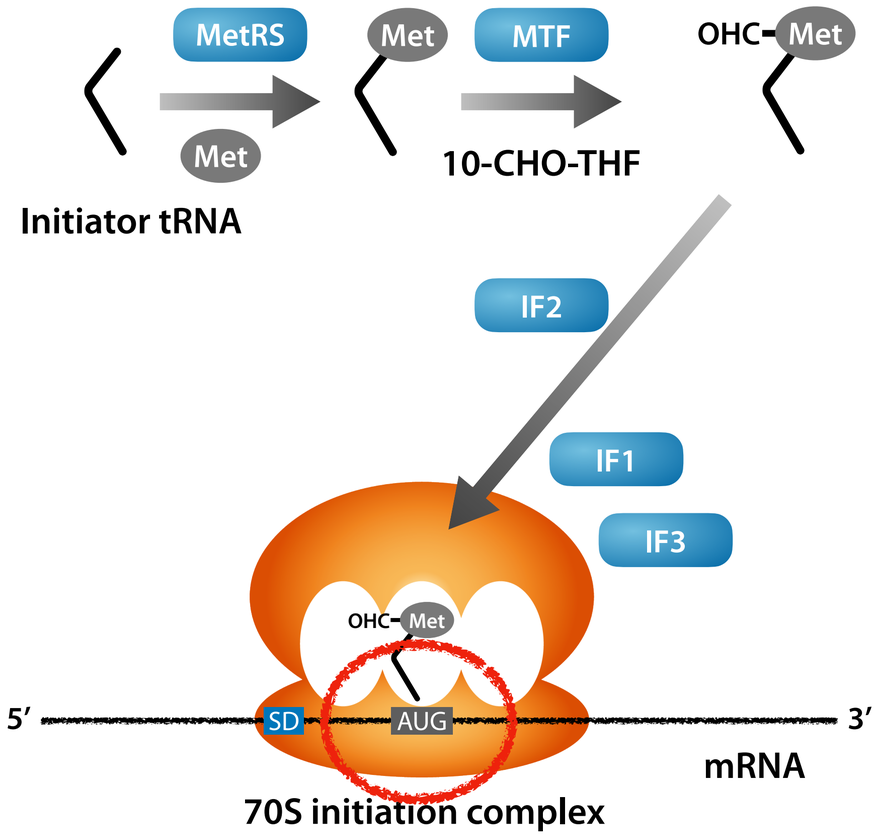
IF1, IF2, IF3



Initiation factors were not always required for the translation reaction.

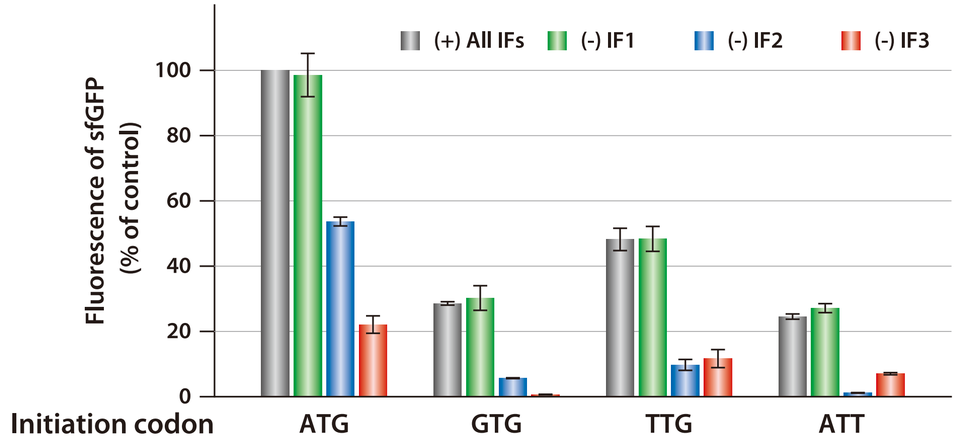
GeneFrontier

Initiation in the PURE system (initiation codon)



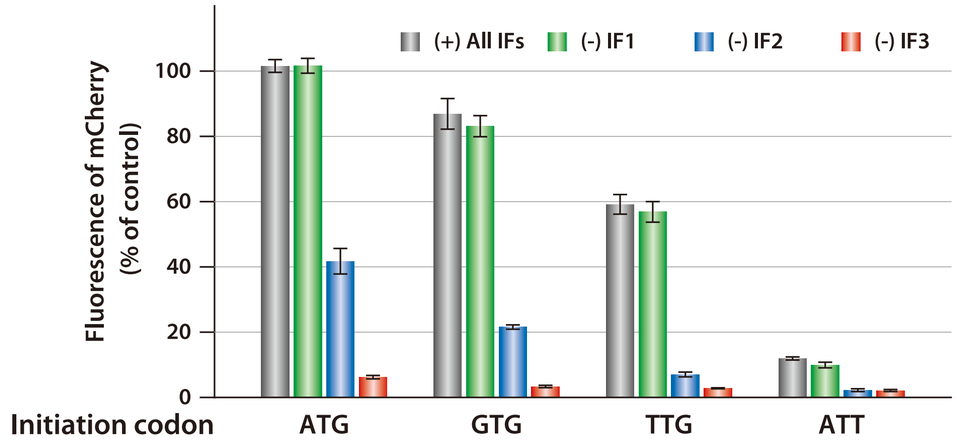
sfGFP

M S K G E E ..
atgtctaaaggtgaagaa..



mCherry

M A I I K E ..
atggctattattaaagaa..



Initiation codons in *E. coli* (Ref; Blattner *et al.* (1997) *Science*, 277, p. 1453)

- AUG; 83%
- GUG; 14% (e.g. EF-Tu, ArgRS, HisRS, RRF)
- UUG; 3%
- AUU; 2 genes (*infC* (IF3), *pcnB*)

GTG, TTG, and ATT functioned as the initiation codon, but the synthesis efficiency was dependent on the target proteins.

GeneFrontier

Summary

Initiation in the PURE system

- All IFs and formylation of initiator methionine are not essential for the translation in the PURE system, but IF2 and IF3 enhance the translation efficiency.

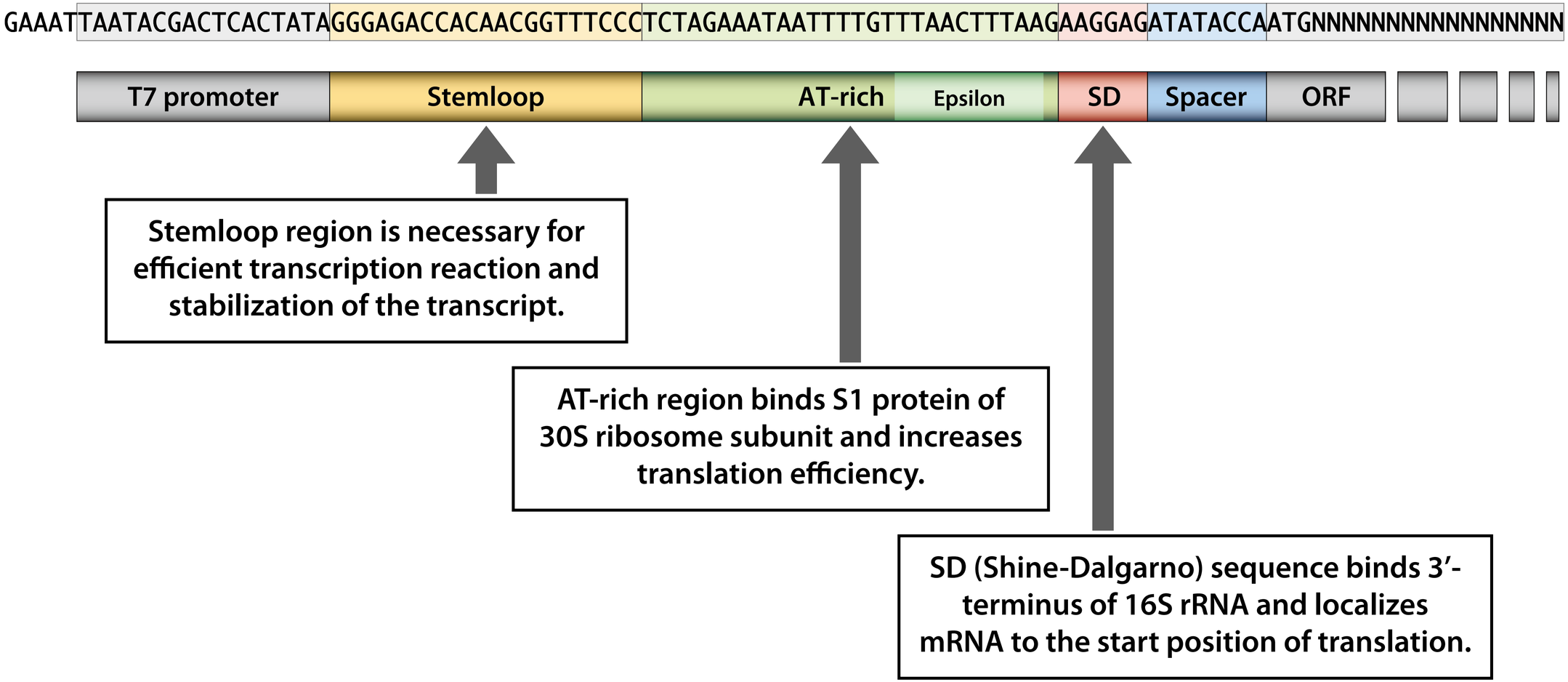
- Codons used as initiation codon in *E. coli* are available for the translation in the PURE system, albeit with low efficiency.

Topics

- ★ PURE system (Translation in *E. coli*)
- ★ Initiation in the PURE system
- ★ **Optimum sequence for the PURE system**
  - ▸ **5’-UTR**
  - ▸ N-terminal region in the ORF

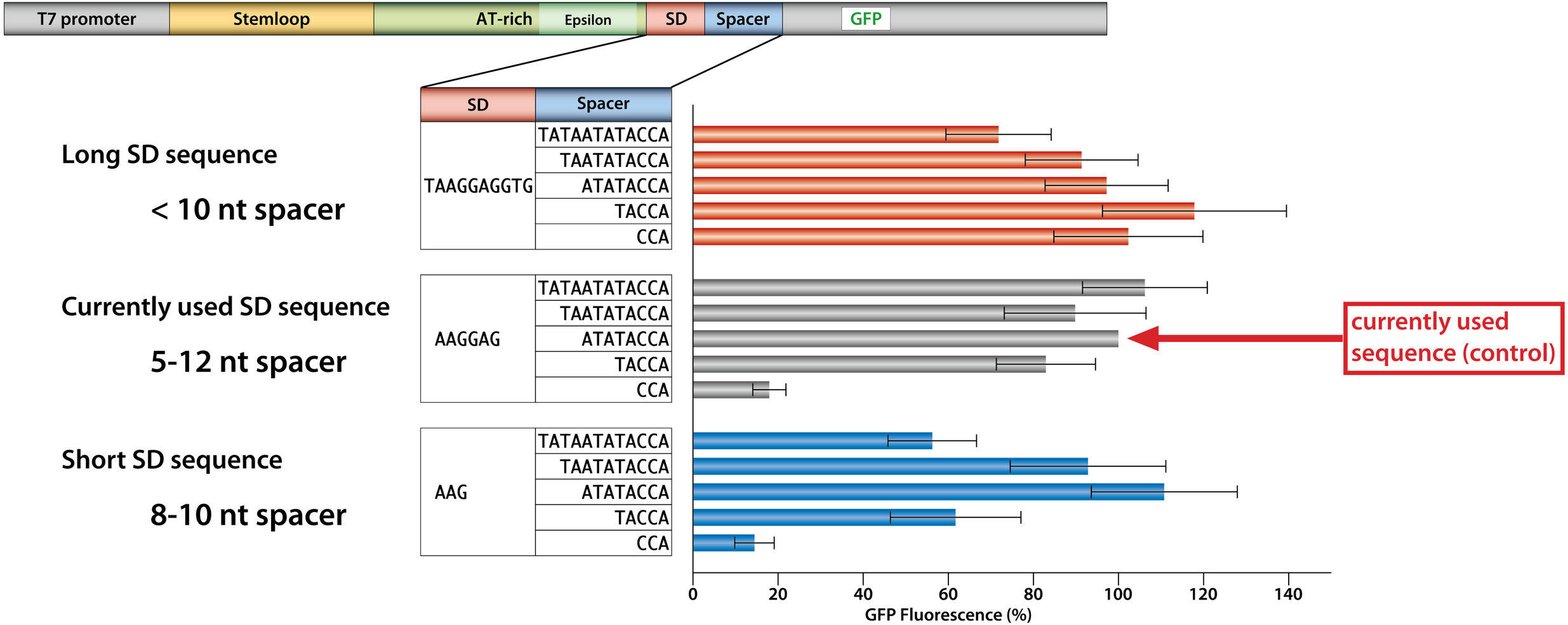
5’ UTR of the template DNA for PURE*frex*

Currently used sequence (derived from T7 *gene 10* UTR)



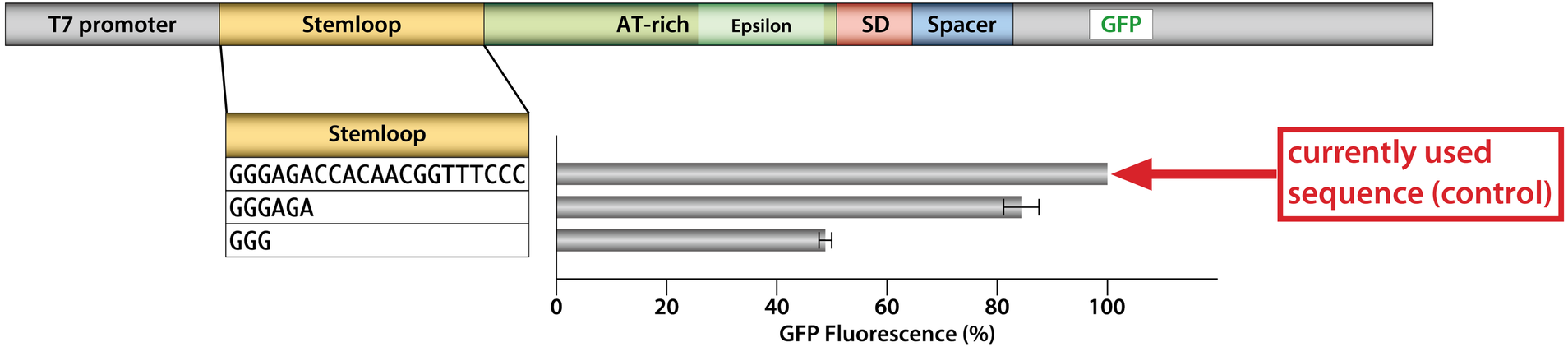
GeneFrontier

SD and Spacer



GeneFrontier

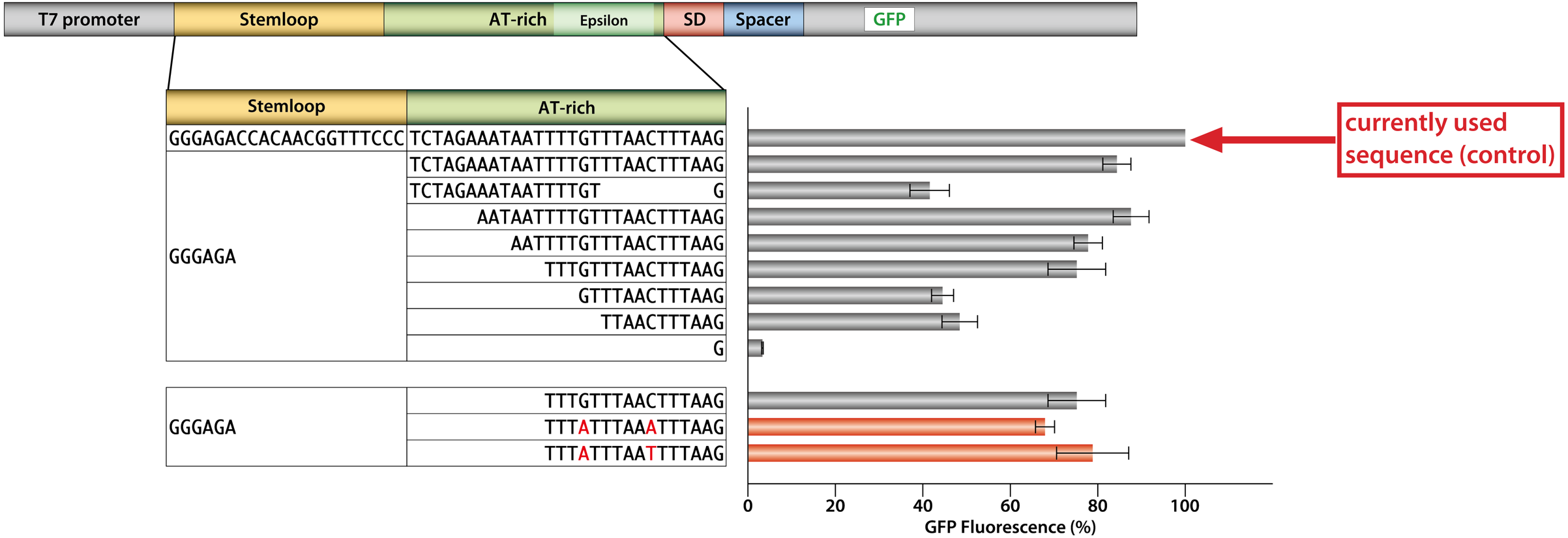
Stemloop



Stemloop region	3 nt	6 nt
	△	○

GeneFrontier

Stemloop and AT-rich



6nt-SL

AT-rich region	No	< 12 nt	> 15 nt
	×	△	○

GeneFrontier

Effect of deletion of 5’ UTR of the downstream gene

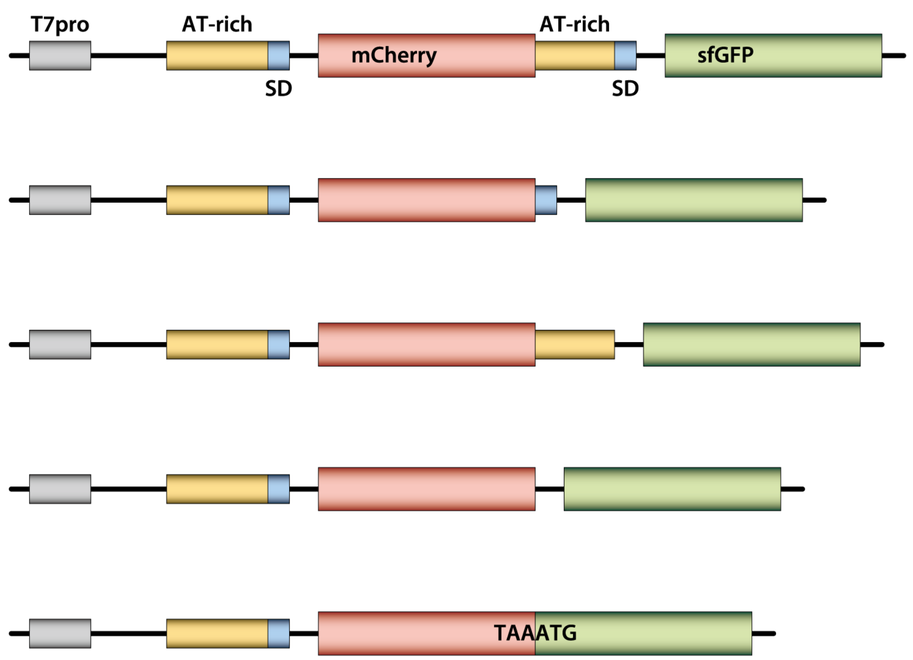
(+) AT-rich/SD

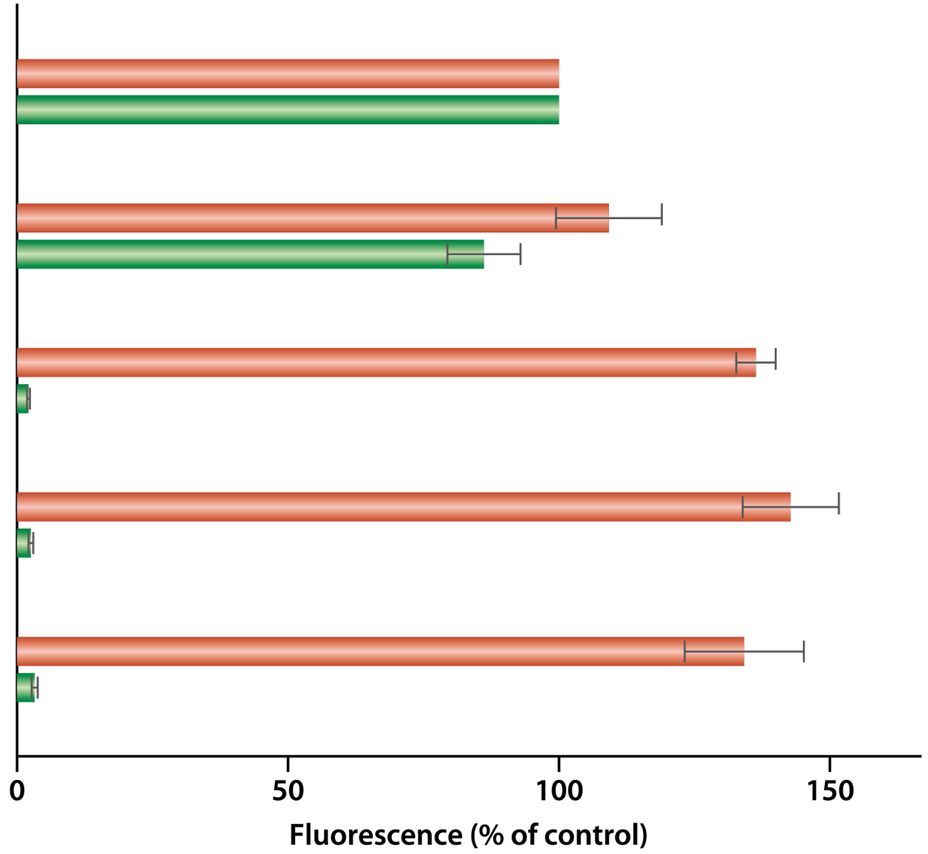
(-) AT-rich

(-) SD

(-) AT-rich/SD

(-) linker

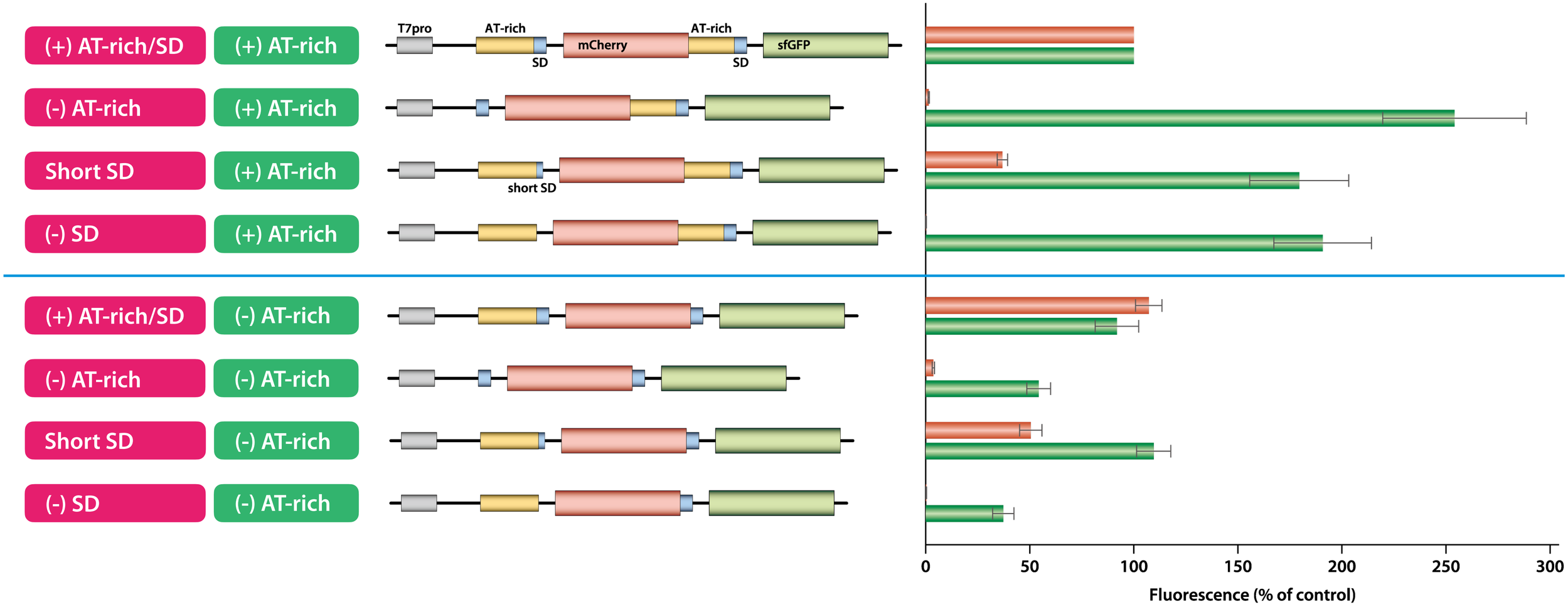




SD sequence was required, but AT-rich region was not, for translation of the downstream gene.

GeneFrontier

Effect of deletion of 5′ UTR of the upstream gene



Deletion of AT-rich region or SD sequence in the upstream gene increased translation of the downstream gene only with AT-rich region in the 5′ UTR.

GeneFrontier

Summary

5' UTR for the translation in the PURE system

Currently used sequence (control)	Efficiency
GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCA	100%
GAAATTAATACGACTCACTATAGGG----------TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCA	50%
GAAATTAATACGACTCACTATAGGGAGA----------TCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCA	85%
GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCC----------GAAGGAGATATACCA	10%
GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAG-----ATATACCA	5%

6 nt for efficient transcription

essential for translation
> 12 nt for efficient translation

essential for translation
> 3 nt for efficient translation

GeneFrontier

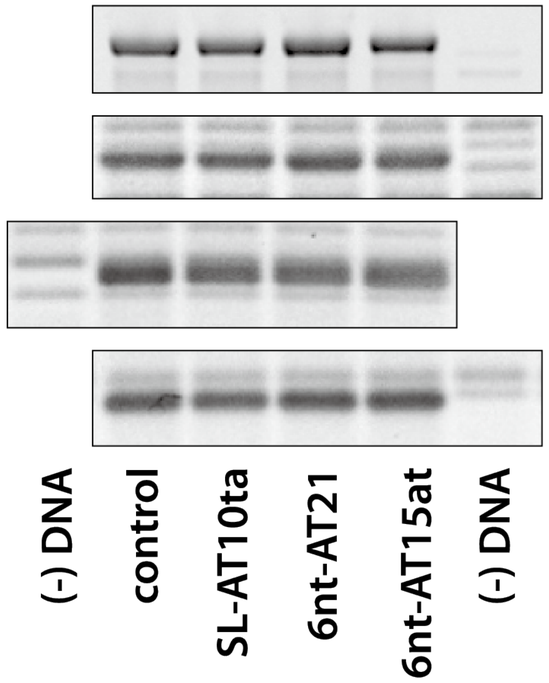
Summary

GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCA **control**

GAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCC----------TTTAACTTTAAAGAAGGAGATATACCA **SL-AT10ta**

GAAATTAATACGACTCACTATAGGGAGA----------AATAATTTTGTTTAACTTTAAGAAGGAGATATACCA **6nt-AT21**

GAAATTAATACGACTCACTATAGGGAGA----------TTTATTTAATTTTAAGAAGGAGATATACCA **6nt-AT15at**



GeneFrontier

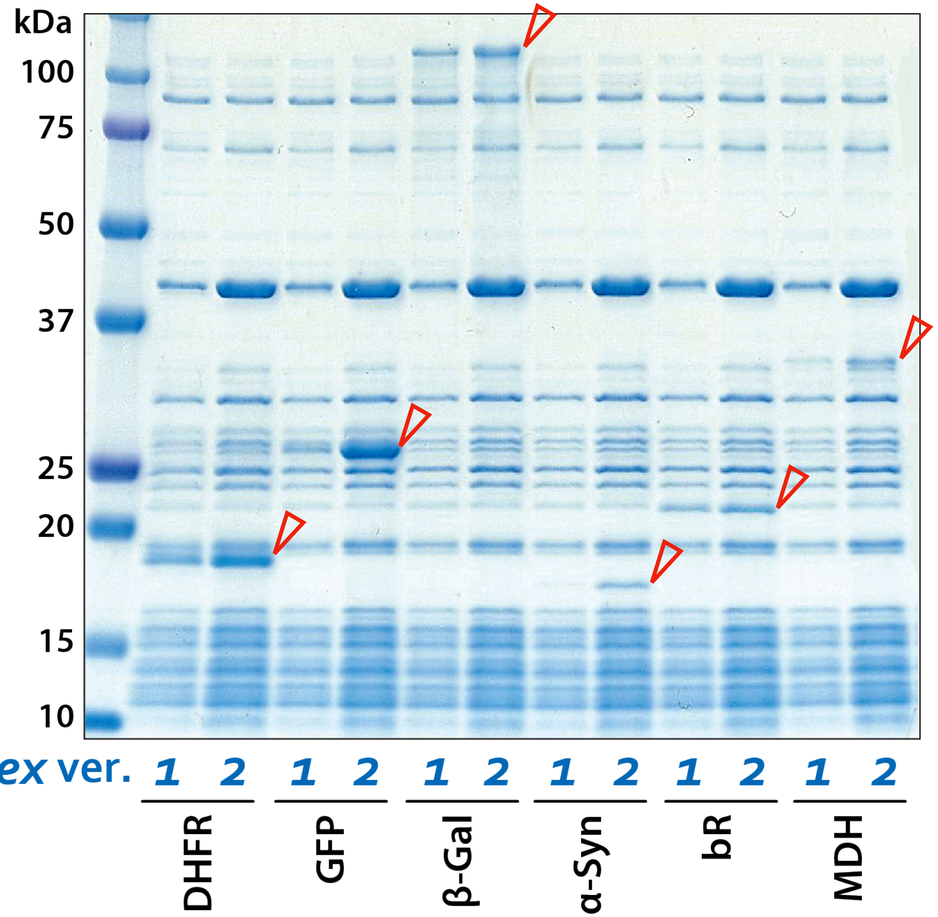
Topics

- ★ PURE system (Translation in *E. coli*)
- ★ Initiation in the PURE system
- ★ **Optimum sequence for the PURE system**
  - ▸ 5’-UTR
  - ▸ **N-terminal region in the ORF**

GeneFrontier

Effect of the sequence on the protein synthesis

Example of protein synthesis



DHFR:
Dihydrofolate reductase
GFP:
Green Fluorescent Protein
β-Gal:
β-Galactosidase
α-Syn:
α-Synuclein
bR:
Bacteriorhodopsin
MDH:
Malate Dehydrogenase

Synthesis efficiency is highly dependent on the target proteins.

GeneFrontier

Effect of N-terminal sequence on the protein synthesis

ARTICLE

https://doi.org/10.1038/s41467-019-13810-1 OPEN

A short translational ramp determines the efficiency of protein synthesis

Manasvi Verma[1,8], Junhong Choi[2,3,6,8], Kyle A. Cottrell[1,8], Zeno Lavagnino[1,7], Erica N. Thomas[4], Slavica Pavlovic-Djuranovic[1], Pawel Szczesny[5], David W. Piston[1], Hani S. Zaher[4], Joseph D. Puglisi[2] & Sergej Djuranovic[1*]

Translation initiation is a major rate-limiting step for protein synthesis. However, recent studies strongly suggest that the efficiency of protein synthesis is additionally regulated by multiple factors that impact the elongation phase. To assess the influence of early elongation on protein synthesis, we employed a library of more than 250,000 reporters combined with in vitro and in vivo protein expression assays. Here we report that the identity of the amino acids encoded by codons 3 to 5 impact protein yield. This effect is independent of tRNA abundance, translation initiation efficiency, or overall mRNA structure. Single-molecule measurements of translation kinetics revealed pausing of the ribosome and aborted protein synthesis on codons 4 and 5 of distinct amino acid and nucleotide compositions. Finally, introduction of preferred sequence motifs only at specific codon positions improves protein synthesis efficiency for recombinant proteins. Collectively, our data underscore the critical role of early elongation events in translational control of gene expression.

Verma *et al.* (2019) *Nat. Commun.*

Peptidyl transferase center decompaction and structural constraints during early protein elongation on the ribosome

Bin Jia[1], Tianlong Wang[1]✉ & Jean Lehmann[2]✉

Peptide bond formation on the ribosome requires that aminoacyl-tRNAs and peptidyl-tRNAs are properly positioned on the A site and the P site of the peptidyl transferase center (PTC) so that nucleophilic attack can occur. Here we analyse some constraints associated with the induced-fit mechanism of the PTC, that promotes this positioning through a compaction around the aminoacyl ester orchestrated by U2506. The physical basis of PTC decompaction, that allows the elongated peptidyl-tRNA to free itself from that state and move to the P site of the PTC, is still unclear. From thermodynamics considerations and an analysis of published ribosome structures, the present work highlights the rational of this mechanism, in which the free-energy released by the new peptide bond is used to kick U2506 away from the reaction center. Furthermore, we show the evidence that decompaction is impaired when the nascent peptide is not yet anchored inside the exit tunnel, which may contribute to explain why the first rounds of elongation are inefficient, an issue that has attracted much interest for about two decades. Results in this field are examined in the light of the present analysis and a physico-chemical correlation in the genetic code, which suggest that elementary constraints associated with the size of the side-chain of the amino acids penalize early elongation events.

Jia *et al.* (2021) *Sci. Rep.*

N-terminal 3rd to 5th amino acid residues are important for efficient translation.

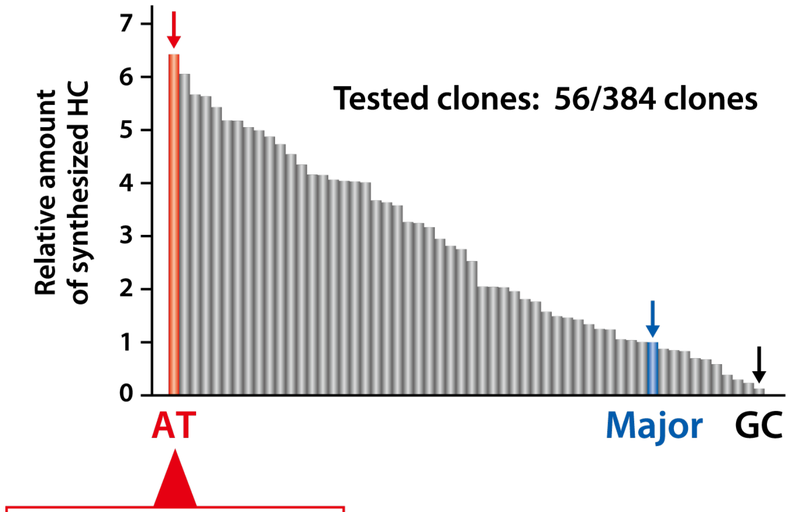
Larger amino acids, which are encoded by AT-rich codon, facilitated early elongation.

GeneFrontier

N-terminal codon suitable for PURE*frex*

N-terminal codon of Trastuzumab HC

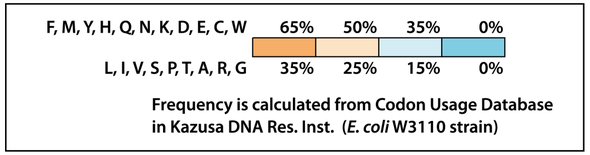
Met	Glu	Val	Gln	Leu	Val
ATG	GAA	GTT	CAA	TTA	GTT
	GAG	GTC	CAG	TTG	GTC
		GTA		CTT	GTA
		GTG		CTC	GTG
				CTA	
				CTG	



Yield: 0.6 mg/mL

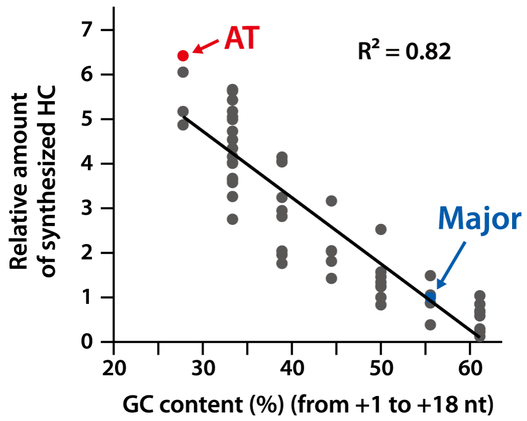
	Met	Glu	Val	Gln	Leu	Val	GC (%)	ΔG (kcal/mol)
AT	ATG	GAA	GTA	CAA	TTA	GTT	28	0.0
Major	ATG	GAA	GTG	CAG	CTG	GTG	56	-4.7
GC	ATG	GAG	GTG	CAG	CTG	GTC	61	-3.9

Codon usage frequency

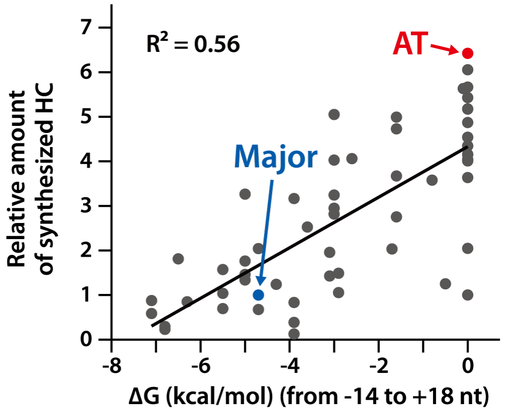


AT-rich codons > Major codons

GC content



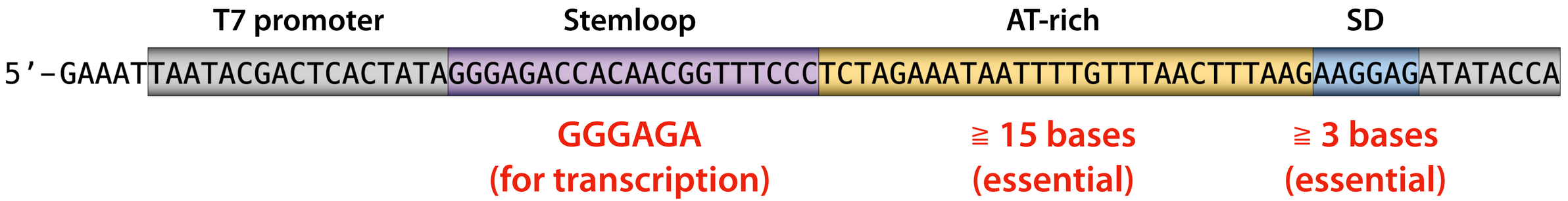
mRNA secondary structure



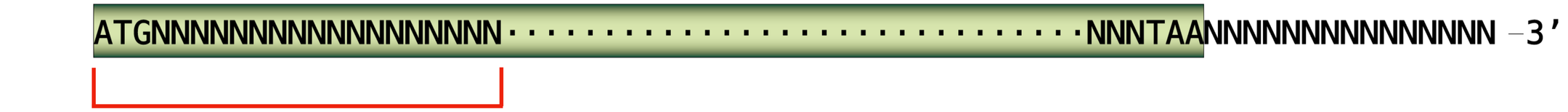
GeneFrontier

Optimum sequence of the template DNA for PURE*frex*

5'UTR (from T7 phage gene10)



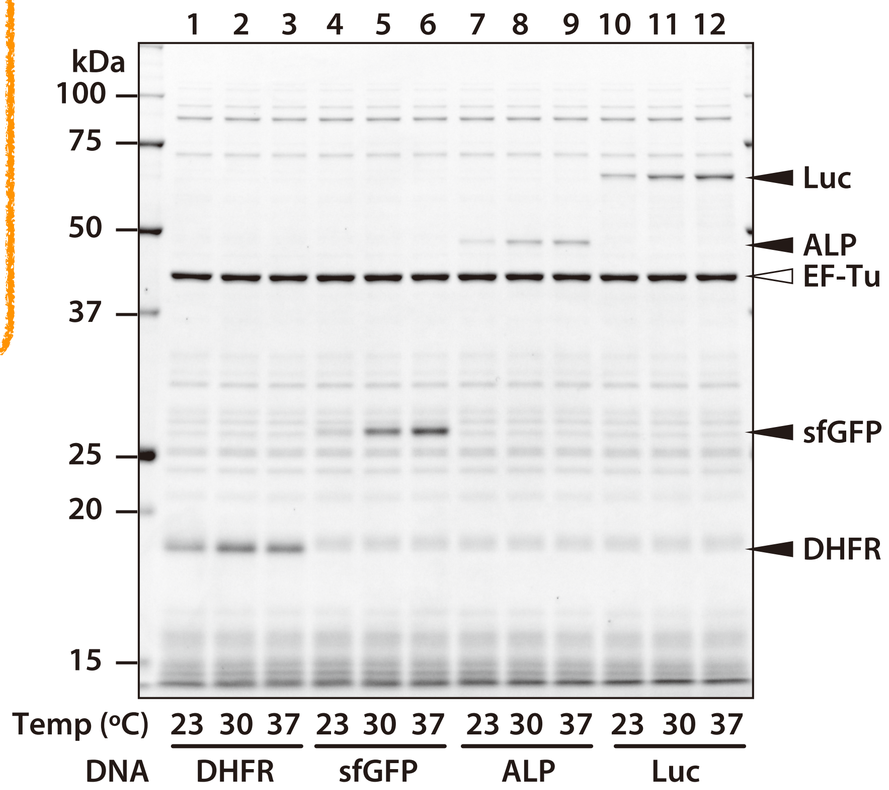
ORF



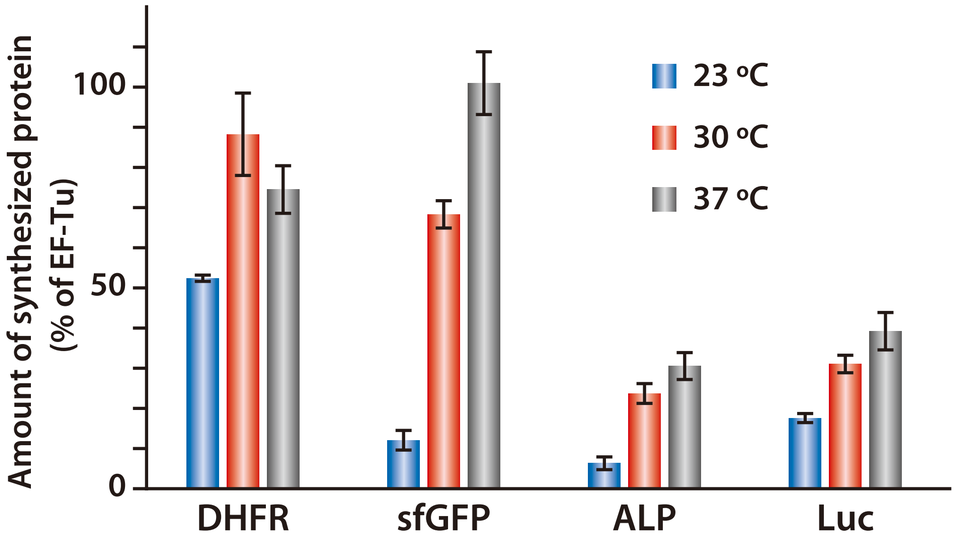
GeneFrontier

Effect of temperature on the protein synthesis

PURE*frex* 2.1 (4 mM GSH)
+ template DNA
incubation
for 24 hours
at 23, 30 or 37°C
SDS-PAGE



Fuse-Murakami *et al.* (2024) *Int. J. Mol. Sci.*, 25, 5264.



Synthesis efficiency at lower temperature depended on the target proteins.

GeneFrontier

Effect of 4th amino acid on sfGFP synthesis

sfGFP

```
M  S  K  G  E  E  L  F  T  G
ATGTCTAAAGGTGAAGAATTATTTACTGGT..
```

	T		C		A		G	
T	TTT	Phe	TCT	Ser	TAT	Tyr	TGT	Cys
	TTC		TCC		TAC		TGC	
	TTA	Leu	TCA		TAA	Stop	TGA	Stop
	TTG		TCG		TAG	Stop	TGG	Trp
C	CTT	Leu	CCT	Pro	CAT	His	CGT	Arg
	CTC		CCC		CAC		CGC	
	CTA		CCA		CAA	Gln	CGA	
	CTG		CCG		CAG		CGG	
A	ATT	Ile	ACT	Thr	AAT	Asn	AGT	Ser
	ATC		ACC		AAC		AGC	
	ATA		ACA		AAA	Lys	AGA	Arg
	ATG	Met	ACG		AAG		AGG	
G	GTT	Val	GCT	Ala	GAT	Asp	GGT	Gly
	GTC		GCC		GAC		GGC	
	GTA		GCA		GAA	Glu	GGA	
	GTG		GCG		GAG		GGG	

GeneFrontier

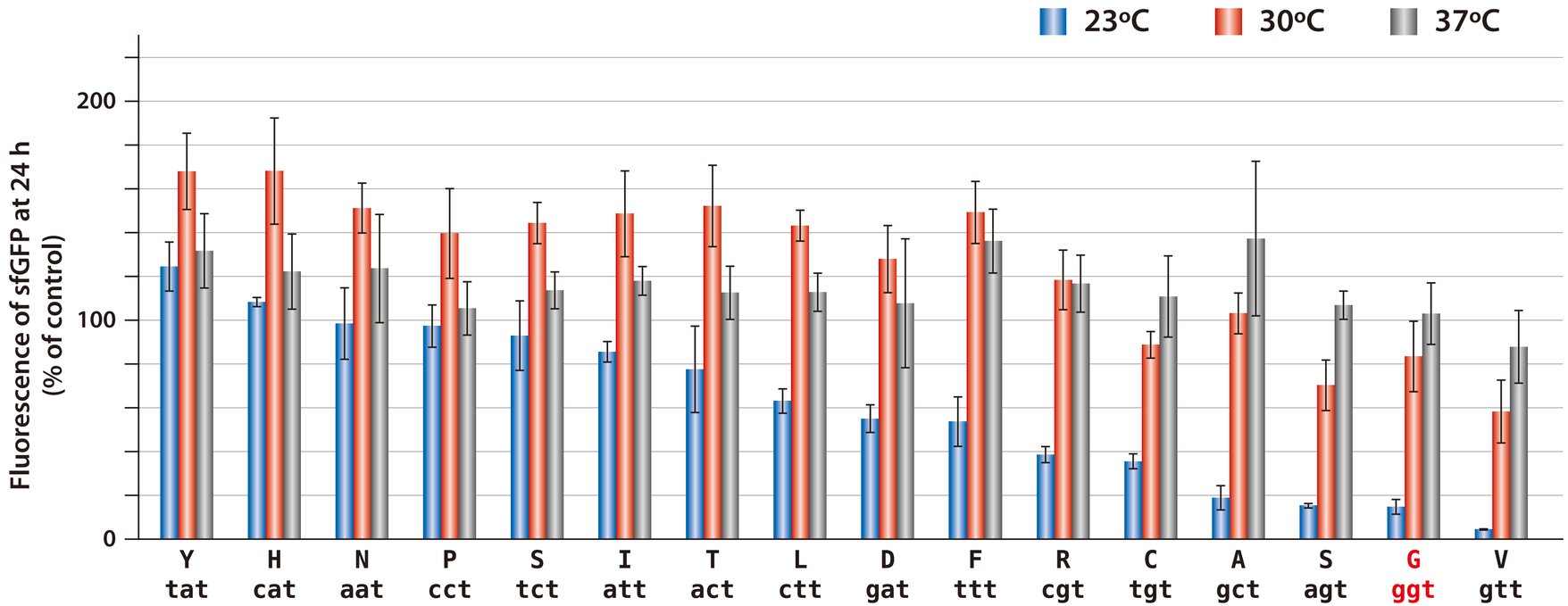
Effect of 4th amino acid on sfGFP synthesis

PURE*frex* 2.1 (4 mM GSH)
+ template DNA
incubation
for 24 hours
at 23, 30 or 37°C

sfGFP

M S K G E E L F T G
ATGTCTAAAGGTGAAGAATTATTTACTGGT..

Fuse-Murakami *et al.* (2024) *Int. J. Mol. Sci.*, 25, 5264.



The fourth amino acid was very important for the efficient translation at lower temperature.

GeneFrontier

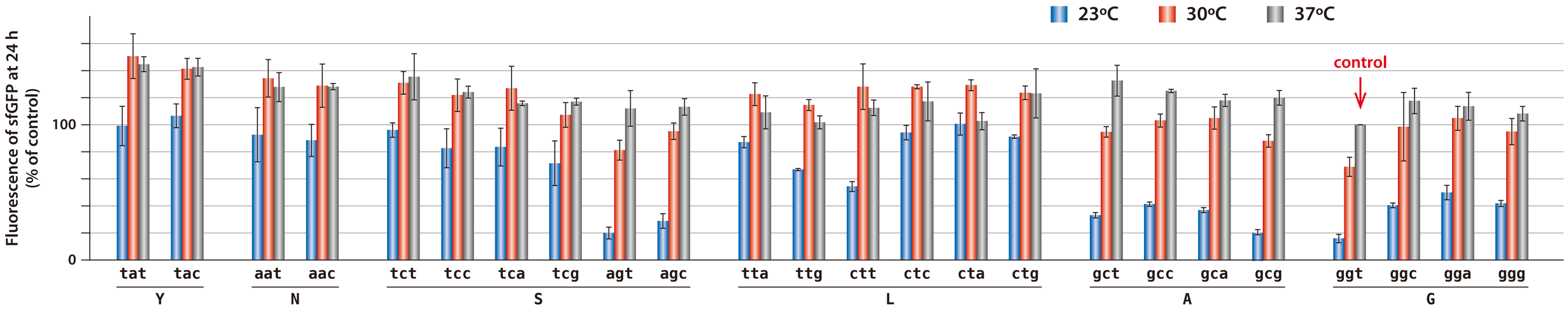
Effect of 4th amino acid on sfGFP synthesis

PURE*frex* 2.1 (4 mM GSH)
+ template DNA
incubation
for 24 hours
at 23, 30 or 37°C

sfGFP

M S K G E E L F T G
ATGTCTAAAGGTGAAGAATTATTTACTGGT..

Fuse-Murakami *et al.* (2024) *Int. J. Mol. Sci.*, 25, 5264.



At the fourth position, the influence of amino acids was greater than that of synonymous codons.

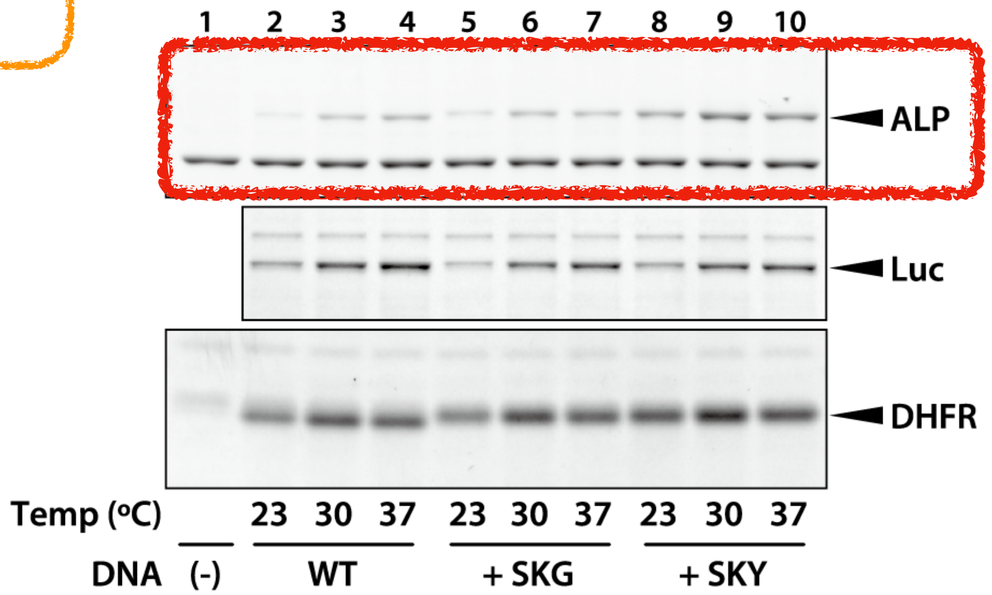
GeneFrontier

Effect of N-terminal tag on the protein synthesis

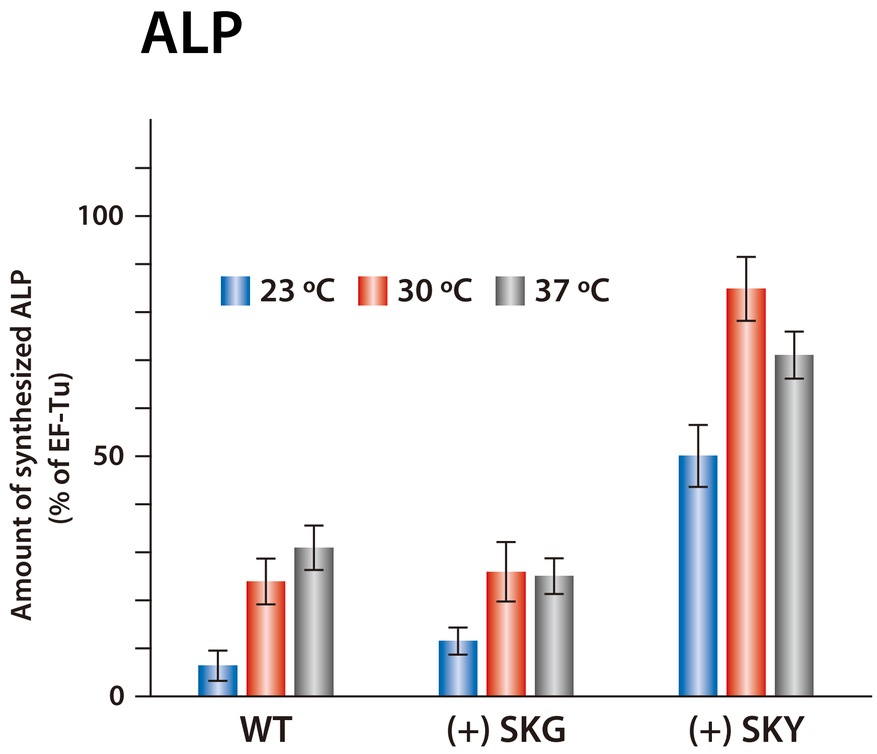
PURE*frex* 2.1 (4 mM GSH)
+ template DNA
incubation
for 24 hours
at 23, 30 or 37°C
SDS-PAGE

```
Luc :  MEDAKNIKKG..
ALP:   MRTPEMPVLE..
DHFR: MISLIAALAV..
```

SKG/SKY



Fuse-Murakami *et al.* (2024) *Int. J. Mol. Sci.*, 25, 5264.



N-terminal additional sequence increased the amount of synthesized ALP at all temperature.

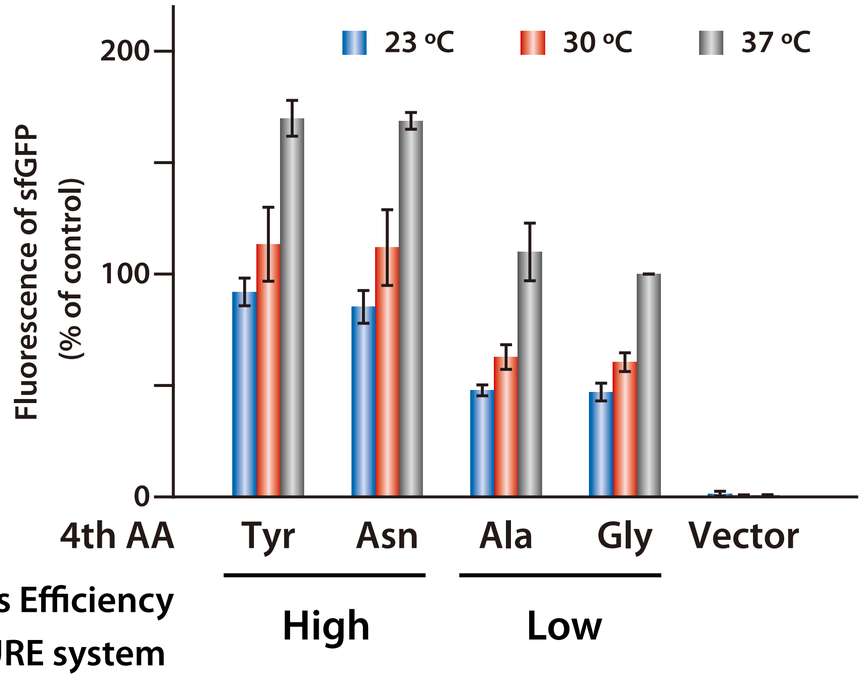
GeneFrontier

Expression of sfGFP variants in *E. coli*

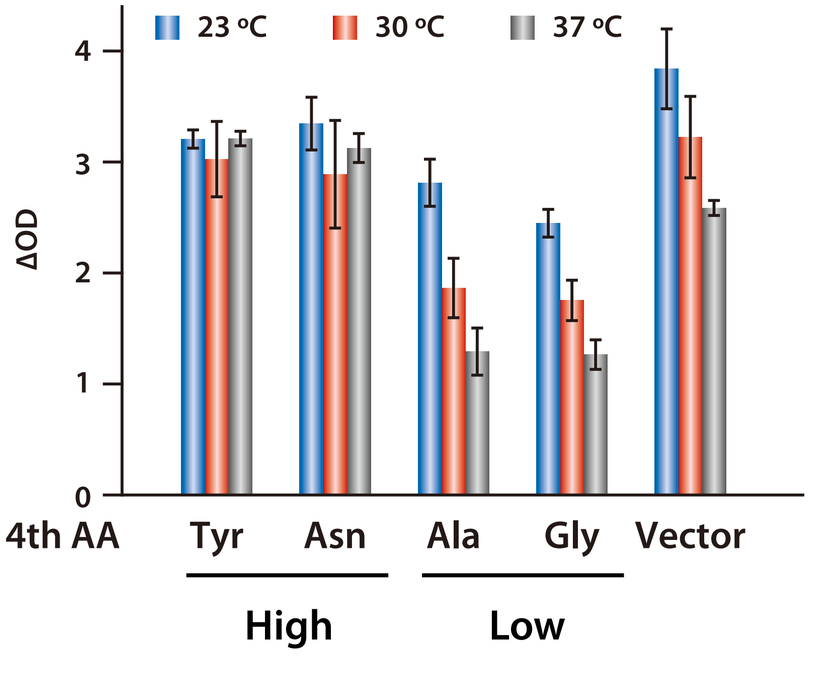
BL21(DE3) + pET-sfGFP DNA
(Tyr/Asn/Ala/Gly@4th)
Incubation at 23, 30 or 37°C
+ IPTG (at OD=1)
Incubation at 23, 30 or 37°C
Measure OD and fluorescence

Fuse-Murakami *et al.* (2024) *Int. J. Mol. Sci.*, 25, 5264.

Fluorescence



ΔOD



Variants that were highly synthesized in the PURE system were also highly expressed in *E. coli* cells.
The difference in the expression level between variants was less than that in the PURE system.
Expression of Ala and Gly variants suppressed the growth of *E. coli* cells.

GeneFrontier

Summary

Optimum sequence for the translation in the PURE system

- AT-rich codons are preferable in the N-terminal region.
- Multiple codons should be used for each amino acid in the entire ORF.
- Amino acids at the N-terminal region are very important for efficient protein synthesis especially at lower temperature.

Summary

Translation in the PURE system = Translation in *E. coli* cells ?

- **Initiation factors are not essential.**
- **AT-rich region in the 5'UTR is important as well as SD sequence.**
- **Translation efficiency at lower temperature is dependent on the sequence of the target proteins.**

Summary

Translation in the PURE system ≈ Translation in *E. coli* cells

- **Initiation factors are not essential.**
- **AT-rich region in the 5'UTR is important as well as SD sequence.**
- **Translation efficiency at lower temperature is dependent on the sequence of the target proteins.**